

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 2004, 19:24:00 ; Search time 120.375 Seconds
(without alignments)
43.019 Million cell updates/sec

Title: US-10-618-644-4
Perfect score: 58
Sequence: 1 IPPGVVWYWT 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 02:.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	517	2	Q7GC77 glycine max
2	58	100.0	517	2	Q39922 glycine soj
3	58	100.0	517	2	Q8SB12 glycine max
4	58	100.0	560	2	Q8S9D0 glycine max
5	58	100.0	562	2	Q43452 glycine max
6	58	100.0	563	2	Q39921 glycine soj
7	58	100.0	563	2	Q9SB11 glycine max
8	57	98.3	136	2	Q43672 vicia faba
9	57	98.3	564	2	Q43673 vicia faba
10	57	98.3	566	2	Q4294 pium sativ
11	56	96.6	536	2	Q9FEC5 glycine max
12	56	96.6	536	2	Q6DR94 glycine max
13	52	89.7	562	1	GLC4_SOYBN
14	49	84.5	308	2	Q43671 vicia faba
15	49	84.5	484	1	LSG4_VICFA
16	49	84.5	485	2	Q41703 vicia sativ
17	49	84.5	503	1	LEGJ_PEA
18	48	82.8	516	1	GLC5_SOYBN
19	48	82.8	517	2	Q93707 glycine max
20	48	82.8	517	2	Q93708 glycine max
21	46	79.3	363	2	Q82427 Oryza sativ
22	44	75.9	165	2	Q8P9R4 xanthomonas
23	44	75.9	427	2	Q8PLJ6 xanthomonas
24	43	74.1	168	2	Q94704 schizosacch
25	43	74.1	183	2	Q6TM56 bacterioph
26	43	74.1	183	2	AAC94484 bacterioph
27	43	74.1	707	2	Q9F348 streptomyc
28	42	72.4	242	2	Q46604 sus scrofa
29	42	72.4	323	2	Q9NUJ7 homo sapien
30	42	72.4	330	2	Q63241 rattus norv
31	42	72.4	331	2	Q63239 rattus norv

32 42 72.4 503 1 NORB_CHLCV Q823p2 chlamydomphi
33 42 72.4 503 1 NORB_CHLMU Q9pkb6 chlamydia m
34 42 72.4 503 1 NORB_CHLPN Q9zbb6 chlamydia p
35 42 72.4 503 1 NORB_CHLTR O84280 chlamydia t
36 42 72.4 509 2 O6MEH4 Q6meh4 parachlamyd
37 42 72.4 509 2 CAF23025 Caf23025 parachlam
38 42 72.4 555 2 Q9VHV2 Q9vrv2 drosophila
39 42 72.4 568 2 Q8GNZ3 Q8gnz3 drosophila
40 42 72.4 705 2 Q63710 Q63710 rattus ratt
41 42 72.4 707 2 Q9TT07 Q9tt07 canis famill
42 42 72.4 723 2 Q86VI4 Q86vi4 homo sapien
43 41.5 71.6 548 2 Q9FJUS Q9fj05 pseudomonas
44 41.5 71.6 561 2 Q9LUN3 Q9lun3 plasmid psb
45 41.5 71.6 569 2 O08449 O08449 pseudomonas

ALIGNMENTS

RESULT 1
Q7GC77 PRELIMINARY; PRT; 517 AA.
AC Q7GC77;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Glycinin A3B4 subunit.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TrISSE=Cotyledon;
RA Adachi M., Katsube T., Masuda T., Utsuni S.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -I- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond (By similarity)
CC -I- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC family.
CC EMBL; AB049440; BAB15802.1; -
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; Rmlc Like cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; 1.
KW Seed storage protein; Storage protein.
SQ SEQUENCE 517 AA; 58187 MW; 0440F4C72181B7C CRC64;

Query Match 100.0%; Score 58; DB 2; Length 517;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPPGVVWYWT 9
Db 149 IPPGVVWYWT 157

RESULT 2
Q39922 PRELIMINARY; PRT; 517 AA.
ID Q39922
AC Q39922;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE G5 protein.
GN Name=G5;
OS Glycine soja (Wild soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

DR PROSITE; PS00305; 11S SEED STORAGE; 1.
KW Seed storage protein; Storage protein.
SQ SEQUENCE 517 AA; 58229 MW; 0E1F51BC72181B7C CRC64;

Query Match 100.0%; Score 58; DB 2; Length 517;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPPGVPTWT 9
Db 149 IPPGVPTWT 157

RESULT 4
ID Q9S9D0 PRELIMINARY; PRT; 560 AA.
AC Q9S9D0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Glycinin G4 subunit.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92393391; PubMed=2485233;
RX Nielsen N.C., Dickinson C.D., Cho T.J., Thanh V.H., Scallan B.J.,
RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;
RT "Characterization of the glycinin gene family in soybean.";
RL Plant Cell 1:313-328(1989).
CC -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond (By similarity).
CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC family.
DR PIR; S11004; S11004.
DR HSP; P04776; 1PXZ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR011051; RmlC_like_cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S SEED STORAGE; 1.
DR PROSITE; PS00305; 11S SEED STORAGE; 1.
SQ SEQUENCE 560 AA; 63679 MW; F95DBEA4012DA024 CRC64;

Query Match 100.0%; Score 58; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IPPGVPTWT 9
Db 149 IPPGVPTWT 157

RESULT 5
ID Q43452 PRELIMINARY; PRT; 562 AA.
AC Q43452;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glycinin precursor.
DE Name=Gy4;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]

```

RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=92256811; PubMed=1316192;
RA Xue Z.T., Xu M.L., Shen W., Zhuang N.L., Hu W.M., Shen S.C.;
RT "Characterization of the Gy4 Glycine gene from soybean Glycine max cv.
RL Forrest";
RN Plant Mol. Biol. 18:897-908(1992).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Hue Z.T.;
RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond (By similarity).
CC -1- SIMILARITY: Belongs to the IIS seed storage protein (globulins)
CC family.
DR EMBL; X52863; CAA37044.1; -.
DR PIR; P00199; P00199.
DR PIR; S20946; S20946.
DR HSP; P04776; IFXZ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR011051; RmlC like_cupin.
DR InterPro; IPR006044; Seedstore_118.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; IIS_SEED_STORAGE; 1.
KW Signal.
FT CHAIN 1 23 Potential.
FT CHAIN 24 562 Glycinin.
SQ SEQUENCE 562 AA; 63876 MW; 3A4EF28B49AB15A CRC64;
Query Match 100.0%; Score 58; DB 2; Length 562;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IPPGVPTWT 9
Db 148 IPPGVPTWT 156
RESULT 6
ID Q39921 PRELIMINARY; PRT; 563 AA.
AC Q39921;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE AS4B3 subunit Gy4;
GN Name=Glycinin Gy4;
OS Glycine soja (Wild soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3848;
[1]
RN EMBL; X86570; CAA60533.1; -.
RP SEQUENCE FROM N.A.
RC STRAIN=SH1;
RA Xue Z.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond (By similarity).
CC -1- SIMILARITY: Belongs to the IIS seed storage protein (globulins)
CC family.
DR EMBL; X86570; CAA60533.1; -.
DR PIR; S54802; S54802.
DR HSP; P04776; IFXZ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR011051; RmlC like_cupin.
DR InterPro; IPR006044; Seedstore_118.

DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; IIS_SEED_STORAGE; 1.
SQ SEQUENCE 563 AA; 63806 MW; EC171BED5F2B361F CRC64;
Query Match 100.0%; Score 58; DB 2; Length 563;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IPPGVPTWT 9
Db 149 IPPGVPTWT 157
RESULT 7
ID Q9SB11 PRELIMINARY; PRT; 563 AA.
AC Q9SB11;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glycinin.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Leaves;
RA Chen S., Arahira M., Fukazawa C.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond (By similarity).
CC -1- SIMILARITY: Belongs to the IIS seed storage protein (globulins)
CC family.
DR EMBL; AB004062; BAA74953.1; -.
DR PIR; P00199; P00199.
DR HSP; P04776; IFXZ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR011051; RmlC like_cupin.
DR InterPro; IPR006044; Seedstore_118.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; IIS_SEED_STORAGE; 1.
SQ SEQUENCE 563 AA; 63797 MW; 4ACC765C55AB9E18 CRC64;
Query Match 100.0%; Score 58; DB 2; Length 563;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IPPGVPTWT 9
Db 149 IPPGVPTWT 157
RESULT 8
ID Q43672 PRELIMINARY; PRT; 136 AA.
AC Q43672;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Legumin; legumin-related high molecular weight polypeptide
DE (Fragment).
GN Name=LelB161;
OS Vicia faba (Broad bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid1; Fabales; Fabaceae; Papilionoideae; Viciaeae; Vicia.
OX NCBI_TaxID=3906;

Sun Nov 7 14:54:04 2004

us-10-618-644-4.rup

```

RN      SEQUENCE FROM N.A.
RC      TISSUE=Cotyledons;
RX      MEDLINE=94272010; PubMed=8003694;
RA      Heim U., Baumlein H., Wobus U.;
RT      "The legumin gene family: a reconstructed Vicia faba legumin gene
RT      encoding a high-molecular-weight subunit is related to type B genes.";
RL      Plant Mol. Biol. 25:131-135(1994).
DR      EMBL; Z26487; CAA81261.1; -.
DR      PIR; T12140; T12140.
DR      GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR      HSSP; P04776; 1FXZ.
DR      InterPro; IPR006045; Cupin.
DR      InterPro; IPR007113; Cupin region.
DR      InterPro; IPR011051; RmlC_like_cupin.
DR      Pfam; PF00190; Cupin; 1.
FT      NON_TER 1
FT      NON_TER 136
SQ      SEQUENCE 136 AA; 15317 MW; 3C24820F2630F069 CRC64;

Query Match      98.3%; Score 57; DB 2; Length 136;
Best Local Similarity 88.9%; Pred. No. 0.26;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IPPGVPIYWT 9
      ||||:||||
Db      50 IPPGIPYWT 58

RESULT 9
ID      Q43673      PRELIMINARY;      PRT;      564 AA.
AC      Q43673;
DT      01-NOV-1996 (T-EMBLrel. 01, Created)
DT      01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT      01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE      Legumin; legumin-related high molecular weight polypeptide
DE      precursor.
GN      Name=LelB3;
OS      Vicia faba (Broad bean).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OX      NCBI_TaxID=3906;
[1]
RN      SEQUENCE FROM N.A.
RC      TISSUE=Cotyledons;
RX      MEDLINE=94272010; PubMed=8003694;
RA      Heim U., Baumlein H., Wobus U.;
RT      "The legumin gene family: a reconstructed Vicia faba legumin gene
RT      encoding a high-molecular-weight subunit is related to type B genes.";
RL      Plant Mol. Biol. 25:131-135(1994).
CC      -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC      basic chain derived from a single precursor and linked by a
CC      disulfide bond (By similarity).
CC      -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC      family.
DR      EMBL; Z26489; CAA81262.1; -.
DR      PIR; S46503; S37241.
DR      HSSP; P04776; 1FXZ.
DR      GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR      InterPro; IPR006045; Cupin.
DR      InterPro; IPR007113; Cupin region.
DR      InterPro; IPR000977; DNA_ligase.
DR      InterPro; IPR011051; RmlC_like_cupin.
DR      InterPro; IPR006044; Seedstore_11s.
DR      Pfam; PF00190; Cupin; 2.
DR      PRINTS; PR00439; 11SGLOBULIN.
DR      PROSITE; PS00305; 11S_SEED_STORAGE; 1.
DR      PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
KW      Seed storage protein; Signal; Storage protein.
FT      SIGNAL 1
FT      SIGNAL 20
SQ      SEQUENCE 564 AA; 64502 MW; 568E497245A1915A CRC64;

[1]
RN      SEQUENCE FROM N.A.
RC      TISSUE=Cotyledons;
RX      MEDLINE=94272010; PubMed=8003694;
RA      Heim U., Baumlein H., Wobus U.;
RT      "The legumin gene family: a reconstructed Vicia faba legumin gene
RT      encoding a high-molecular-weight subunit is related to type B genes.";
RL      Plant Mol. Biol. 25:131-135(1994).
CC      -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC      basic chain derived from a single precursor and linked by a
CC      disulfide bond (By similarity).
CC      -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC      family.
DR      EMBL; Z26489; CAA81262.1; -.
DR      PIR; S46503; S37241.
DR      HSSP; P04776; 1FXZ.
DR      GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR      InterPro; IPR006045; Cupin.
DR      InterPro; IPR007113; Cupin region.
DR      InterPro; IPR000977; DNA_ligase.
DR      InterPro; IPR011051; RmlC_like_cupin.
DR      InterPro; IPR006044; Seedstore_11s.
DR      Pfam; PF00190; Cupin; 2.
DR      PRINTS; PR00439; 11SGLOBULIN.
DR      PROSITE; PS00305; 11S_SEED_STORAGE; 1.
DR      PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
KW      Seed storage protein; Signal; Storage protein.
FT      SIGNAL 1
FT      SIGNAL 20
SQ      SEQUENCE 564 AA; 64502 MW; 568E497245A1915A CRC64;

Query Match      98.3%; Score 57; DB 2; Length 564;
Best Local Similarity 88.9%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IPPGVPIYWT 9
      ||||:||||
Db      142 IPPGIPYWT 150

RESULT 10
ID      Q24294      PRELIMINARY;      PRT;      566 AA.
AC      Q24294;
DT      01-JAN-1998 (T-EMBLrel. 05, Created)
DT      01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT      01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE      Legumin (Minor small) precursor.
GN      Name=LegS;
OS      Pisum sativum (Garden pea).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX      NCBI_TaxID=3888;
[1]
RN      SEQUENCE FROM N.A.
RC      TISSUE=Cotyledon;
RA      Bown D.O., Gatehouse J.A.;
RL      Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
[2]
RN      SEQUENCE FROM N.A.
RC      TISSUE=Cotyledon;
RA      Bown D.P.;
RL      Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
CC      -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC      basic chain derived from a single precursor and linked by a
CC      disulfide bond (By similarity).
CC      -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC      family.
DR      EMBL; X67424; CAA47809.1; -.
DR      PIR; T06453; T06453.
DR      HSSP; P04776; 1FXZ.
DR      GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR      InterPro; IPR006045; Cupin.
DR      InterPro; IPR007113; Cupin region.
DR      InterPro; IPR000977; DNA_ligase.
DR      InterPro; IPR011051; RmlC_like_cupin.
DR      InterPro; IPR006044; Seedstore_11s.
DR      Pfam; PF00190; Cupin; 2.
DR      PRINTS; PR00439; 11SGLOBULIN.
DR      PROSITE; PS00305; 11S_SEED_STORAGE; 1.
DR      PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
KW      Seed storage protein; Signal; Storage protein.
FT      SIGNAL 1
FT      SIGNAL 22
SQ      SEQUENCE 566 AA; 64872 MW; 663A80DE265359CE CRC64;

Query Match      98.3%; Score 57; DB 2; Length 566;
Best Local Similarity 88.9%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 IPPGVPIYWT 9
      ||||:||||
Db      144 IPPGIPYWT 152

RESULT 11
ID      Q9FEC5      PRELIMINARY;      PRT;      536 AA.
AC      Q9FEC5;
DT      01-MAR-2001 (T-EMBLrel. 16, Created)
DT      01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT      05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE      Glycinin subunit G7.
FT      SIGNAL 1
FT      SIGNAL 20
SQ      SEQUENCE 536 AA; 64502 MW; 568E497245A1915A CRC64;

```

```

GN Name=GY7;
OS Glycine max (Soybean);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Beilinson V., Chen Z., Shoemaker R.C., Fischer R.L., Goldberg R.B.,
RA Nielsen N.C.;
RT "Genomic organization of glycinin genes in soybean.";
RL Theor. Appl. Genet. 104:1132-1140 (2002).
DR EMBL; AF319777; AAG42489.1; -.
DR EMBL; AF319776; AAG42488.1; -.
DR HSSP; P04776; IPXZ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR SEQUENCE 536 AA; 60486 MW; 40F452FAA067FBC7 CRC64;
SQ
Query Match 96.6%; Score 56; DB 2; Length 536;
Best Local Similarity 77.8%; Pred. No. 1.5;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IPFGVPYWT 9
DB 137 VPGGIPYWT 145
:|||||

RESULT 12
Q6DR94 PRELIMINARY; PRT; 536 AA.
AC Q6DR94
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Glycinin subunit G7.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Mi D., Li J.Y., Zhang W.;
RC TISSUE=Leaf;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY649099; AAT68239.1; -.
DR SEQUENCE 536 AA; 60514 MW; 94EF37198D1C4468 CRC64;
SQ
Query Match 96.6%; Score 56; DB 2; Length 536;
Best Local Similarity 77.8%; Pred. No. 1.5;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IPFGVPYWT 9
DB 137 VPGGIPYWT 145
:|||||

RESULT 13
GLC4 SOYBN STANDARD; PRT; 562 AA.
AC P02858;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Glycinin G4 precursor [Contains: Glycinin A5 subunit; Glycinin A4 subunit; Glycinin B3 subunit].
GN Name=GY4;

```

```

OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Bonminori;
RX MEDLINE=8530642; PubMed=2988947;
RA Momma T., Negoro T., Hirano H., Uchimoto A., Uchimoto K., Fukazawa C.;
RT "Glycinin A5A4B3 mRNA: cDNA cloning and nucleotide sequencing of a
RT splitting storage protein subunit of soybean.";
RL Eur. J. Biochem. 149:491-496 (1985).
RN [2]
RP SEQUENCE (A4/A5 SUBUNITS).
RA STRAIN=cv. Bonminori;
RX Hirano H., Fukazawa C., Harada K.;
RT "The primary structures of the A4 and A5 subunits are highly
RT homologous to that of the A3 subunit in the glycinin seed storage
RT protein of soybean.";
RL FEBS Lett. 181:124-128 (1985).
RN [3]
RP SEQUENCE OF 181-386 FROM N.A.
RA STRAIN=cv. CX635-1-1-1;
RX Scallan B.J., Dickinson C.D., Nielsen N.C.;
RT "Characterization of a null-allele for the G4 glycinin gene from
RT soybean.";
RL Mol. Gen. Genet. 208:107-113 (1987).
CC -!- FUNCTION: Glycinin is the major seed storage protein of soybean.
CC -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond.
CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X02626; CAA26478.1; -.
CC EMBL; X05652; CAB57802.1; -.
CC PIR; A91145; FWSYGS.
CC HSSP; P04776; IPXZ.
CC InterPro; IPR006045; Cupin.
CC InterPro; IPR011051; RmlC_like_cupin.
CC InterPro; IPR006044; Seedstore_11s.
CC Pfam; PF00190; Cupin; 2.
CC PRINTS; PR00439; 11SGLOBULIN.
CC PROSITE; PS00305; 11S_SEED_STORAGE; 1.
KW Direct protein sequencing; Multigene family; Seed storage protein;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 120 Glycinin A5 subunit.
FT CHAIN 121 377 Glycinin A4 subunit.
FT CHAIN 378 562 Glycinin B3 subunit.
FT DISULFID 108 384 Interchain (By similarity).
FT CONFLICT 29 29 L -> F (in Ref. 2).
FT CONFLICT 82 82 S -> L (in Ref. 2).
FT CONFLICT 86 86 S -> L (in Ref. 2).
FT CONFLICT 94 94 I -> V (in Ref. 2).
FT CONFLICT 101 103 LGV -> IGM (in Ref. 2).
FT CONFLICT 105 105 I -> F (in Ref. 2).
FT CONFLICT 117 117 E -> Q (in Ref. 2).
FT CONFLICT 253 253 E -> Q (in Ref. 3).
FT CONFLICT 332 335 NKTG -> EODQ (in Ref. 3).
SQ SEQUENCE 562 AA; 63587 MW; F5A06B8856B9BBD6 CRC64;
Query Match 89.7%; Score 52; DB 1; Length 562;
Best Local Similarity 88.9%; Pred. No. 7;

```

CC	acids in seed meals.
CC	- - SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC	basic chain derived from a single precursor and linked by a
CC	disulfide bond
CC	- - SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC	family.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; X03677; CAA27313.1; -
DR	HSSP; P04776; 1FXZ
DR	InterPro; IPR006045; Cupin.
DR	InterPro; IPR007113; Cupin_region.
DR	InterPro; IPR011051; RmlC_like_cupin.
DR	InterPro; IPR006044; Seedstore_11s.
DR	Pfam; PF00190; Cupin; 2.
DR	PRINTS; PR00439; 11SGLOBULIN.
DR	PROSITE; PS00305; 11S_SEED_STORAGE; 1.
KW	Multigene family; Seed storage protein; Signal.
FT	SIGNAL 1 22 Potential.
FT	CHAIN 23 303 Legumin type B alpha chain.
FT	CHAIN 304 484 Legumin type B beta chain.
FT	DISULFID 109 310 Interchain (alpha-beta) (potential).
FT	DOMAIN 111 134 Gln-rich.
FT	DOMAIN 194 230 Gln/Glu-rich.
FT	DOMAIN 284 302 Gln/Glu-rich.
SQ	SEQUENCE 484 AA; 54447 MW; 3A40F53F43FD737 CRC64;
	Query Match 84.5%; Score 49; DB 1; Length 484;
	Best Local Similarity 77.8%; Pred. No. 18;
	Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY	1 IPPGVPIYWT 9
	:
DB	147 IPSGIPYWT 155
	Search completed: November 6, 2004, 19:52:53
	Job time : 122.375 secs

Matches	8;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;																					
QY	1	IPPGVPYWT	9																											
Db	149	IPPSVPYWT	157																											
RESULT 14																														
Q43671	PRELIMINARY; PRT; 308 AA.																													
AC	Q43671;																													
DT	01-NOV-1996	(TREMBLrel. 01, Created)																												
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)																												
DT	01-JUN-2003	(TREMBLrel. 24, Last annotation update)																												
DE	Storage protein.																													
OS	Vicia faba	(Broad bean).																												
OC	Eukaryota;	Viridiplantae;	Streptophyta;	Embryophyta;	Tracheophyta;																									
OC	Spermatophyta;	Magnoliophyta;	eudicotyledons;	core eudicots;	rosids;																									
OC	eurosid1;	Fabales;	Fabaceae;	Papilionoideae;	Viciae;	Vicia.																								
OX	NCBI_TaxID=3906;																													
RN	[1]																													
RP	SEQUENCE FROM N.A.																													
RC	TISSUE=Leaf;																													
RX	MEDLINE=91370849;	PubMed=2491681;																												
RA	Heim U., Schubert R., Baumlein H., Wobus U.;																													
RT	"The legumin gene family: structure and evolutionary implications of																													
RT	Vicia faba B-type genes and pseudogenes.";																													
RL	Plant Mol. Biol. 13:653-663(1989).																													
DR	EMBL; X14238;	CAA32455.1; -																												
DR	HSSP; P04776;	1FXZ.																												
GO	GO:0045735;	Nutrient reservoir activity; IEA.																												
DR	InterPro; IPR006045;	Cupin.																												
DR	InterPro; IPR007113;	Cupin_region.																												
DR	InterPro; IPR011051;	RmlC_like_cupin.																												
DR	InterPro; IPR006044;	Seedstore_11s.																												
DR	Pfam; PF00190;	Cupin; 2.																												
DR	PRINTS; PR00439;	11SGLOBULIN.																												
SQ	SEQUENCE	308 AA; 34233 MW;	E6CD2B1B1D063B99	CRC64;																										
Query Match 84.5%; Score 49; DB 2; Length 308;																														
Best Local Similarity 77.8%; Pred. No. 11;																														
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;																														
QY	1	IPPGVPYWT	9																											
Db	147	IPSGIPYWT	155																											
RESULT 15																														
LEG4_VICFA	STANDARD; PRT; 484 AA.																													
ID	LEG4_VICFA																													
AC	P05130;																													
DT	13-AUG-1987	(Rel. 05, Created)																												
DT	13-AUG-1987	(Rel. 05, Last sequence update)																												
DT	05-JUL-2004	(Rel. 44, Last annotation update)																												
DE	Legumin type B precursor [Contains: Legumin type B alpha chain																													
DE	(Legumin type B acidic chain); Legumin type B beta chain (Legumin type																													
DE	B basic chain)].																													
GN	Name=LEG4;																													
OS	Vicia faba	(Broad bean).																												
OC	Eukaryota;	Viridiplantae;	Streptophyta;	Embryophyta;	Tracheophyta;																									
OC	Spermatophyta;	Magnoliophyta;	eudicotyledons;	core eudicots;	rosids;																									
OC	eurosid1;	Fabales;	Fabaceae;	Papilionoideae;	Viciae;	Vicia.																								
OX	NCBI_TaxID=3906;																													
RN	[1]																													
RP	SEQUENCE FROM N.A.																													
RX	MEDLINE=86176760;	PubMed=3960730;																												
RA	Baumlein H., Wobus U., Pustell J., Kafatos F.C.;																													
RT	"The legumin gene family: structure of a B type gene of Vicia faba and																													
RT	a possible legumin gene specific regulatory element.";																													
RL	Nucleic Acids Res. 14:2707-2720(1986).																													
CC	- -	FUNCTION: This protein found in the seeds of many leguminous and																												
CC	nonleguminous plants is the source of sulfur-containing amino																													

acids in seed meals.

-|- SUBUNIT: Hexamer; each subunit is composed of an acidic and a basic chain derived from a single precursor and linked by a disulfide bond.

-|- SIMILARITY: Belongs to the 11S seed storage protein (globulins) family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X03677; CAA27313.1; --

HSSP; P04776; 1FXZ.

InterPro; IPR006045; Cupin.

InterPro; IPR007113; Cupin_region.

InterPro; IPR011051; RmlC_like_cupin.

InterPro; IPR006044; Seedstore_11s.

Pfam; PF00190; Cupin; 2.

PRINTS; PR00439; 11SGLOBULIN.

PROSITE; PS00305; 11S_SEED_STORAGE; 1.

Multigene family; Seed storage protein; Signal.

SIGNAL 1 22 Potential.

CHAIN 23 303 Legumin type B alpha chain.

CHAIN 304 484 Legumin type B beta chain.

DISULFID 109 310 Interchain (alpha-beta) (Potential).

DOMAIN 111 134 Gln-rich.

DOMAIN 194 230 Gln/Glu-rich.

DOMAIN 284 302 Gln/Glu-rich.

SEQUENCE 484 AA; 54447 MW; 3A40F53F43F3D737 CRC64;

Query Match 84.5%; Score 49; DB 1; Length 484;

Best Local Similarity 77.8%; Pred. No. 18;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IPPGVPYWT 9

Db 147 IPSGIPYWT 155

Search completed: November 6, 2004, 19:52:53

Job time : 122.375 secs